

Constraints on Alternatives to Supermassive Black Holes

M. Coleman Miller¹

E-mail: miller@astro.umd.edu

¹ *University of Maryland, Department of Astronomy, College Park, MD 20742-2421, USA;
and Goddard Space Flight Center, Greenbelt, MD, USA*

Submitted 2005 October

ABSTRACT

Observations of the centers of galaxies continue to evolve, and it is useful to take a fresh look at the constraints that exist on alternatives to supermassive black holes at their centers. We discuss constraints complementary to those of Maoz (1998) and demonstrate that an extremely wide range of other possibilities can be excluded. In particular, we present the new argument that for the velocity dispersions inferred for many galactic nuclei, even binaries made of point masses cannot stave off core collapse because hard binaries are so tight that they merge via emission of gravitational radiation before they can engage in three-body or four-body interactions. We also show that under these conditions core collapse leads inevitably to runaway growth of a central black hole with a significant fraction of the initial mass, regardless of the masses of the individual stars. For clusters of noninteracting low-mass objects (from low-mass stars to elementary particles), relaxation of stars and compact objects that pass inside the dark region will be accelerated by interactions with the dark mass. If the dark region is instead a self-supported object such as a fermion ball, then if stellar-mass black holes exist they will collide with the object, settle, and consume it. The net result is that the keyhole through which alternatives to supermassive black holes must pass is substantially smaller and more contrived than it was even a few years ago.

Key words:

black hole physics — galaxies: kinematics and dynamics — Galaxy: center — Galaxy: nucleus — gravitation

1 INTRODUCTION

High-resolution observations of the nuclei of many galaxies have revealed large dark masses in small regions. These are most naturally interpreted as supermassive black holes, but as emphasized by Maoz (1998) it is important to take stock of how rigorously we can rule out other possibilities.

Here we present arguments showing that under extremely general conditions almost all other options are ruled out, further emphasizing that supermassive black holes are by far the least exotic and most reasonable explanations for the data in many specific sources. In § 2 we lay out our assumptions, making them as conservative as possible so that our conclusions are robust. In § 3 we show that for many observed galactic nuclei, binaries are unable to heat the stellar distribution effectively because if they are hard then they merge quickly via gravitational radiation. This important constraint, which depends only on dynamics and not the detailed properties of the specific objects, was not presented by Maoz (1998) or elsewhere as far as we are aware. In § 4 we explore the consequences of core collapse and demon-

strate that a very significant mass will inevitably coalesce even for point masses. In § 5 we investigate for the first time the consequences if stellar-mass black holes exist outside the nucleus. We show that enough of them will find their way to the center that they will have serious effects on the nuclear region, likely consuming a significant amount of mass and leading to a supermassive black hole. We discuss the consequences of this analysis in § 6.

2 ASSUMPTIONS AND DYNAMICS

In the spirit of Maoz (1998), we make a series of conservative assumptions to rule out alternatives to supermassive black holes. Let us suppose that observations have revealed that a mass M is confined within a spherically symmetric region whose radius is at most R . We also assume that this mass is composed of identical point masses m ; the point mass assumption minimises the interaction between the masses, and making them identical increases as much as possible the re-

laxation time, on which the masses concentrate in the center of the distribution and hence increase interaction rates. The local two-body relaxation time for a mass m in a region of mass density ρ and velocity dispersion σ is (Spitzer 1987)

$$t_{\text{rlx}} \approx \frac{1}{3 \ln \Lambda} \frac{\sigma^3}{G^2 m \rho} \quad (1)$$

where $\ln \Lambda \sim 10$ is the Coulomb logarithm. In general this time depends on radius, but note that if $\rho \sim r^{-3/2}$ and the velocity dispersion is dominated by a single large mass, the relaxation time is constant with radius.

Any streaming motion (e.g., rotation) reduces the relative speed σ and hence reduces the relaxation time (see, e.g., Kim, Lee, & Spurzem 2004 for a numerical treatment of a rotating stellar system). Therefore, completely random motion leads to the largest timescales.

For N identical masses in a region whose crossing time is t_{cross} , the global relaxation time is approximately (Binney & Tremaine 1987)

$$t_{\text{rlx}} \approx \frac{0.14N}{\ln(0.4N)} t_{\text{cross}} \quad (2)$$

$$\approx 10^9 \text{ yr} M_8^{1/2} (1 M_\odot/m) (R/1 \text{ pc})^{3/2}$$

where $M = 10^8 M_8 M_\odot$. Both expressions for the relaxation time show that for fixed mass density, lower-mass objects take longer to alter their distribution, as is expected because two-body relaxation occurs due to graininess of the gravitational potential, which is less when there are more objects.

The more concentrated the initial density distribution is, the shorter will be the central relaxation time (see the extensive discussion in Quinlan 1996b). To be conservative, we therefore assume a relatively flat distribution such as a Plummer sphere, in which $\rho \propto (1 + r^2/r_c^2)^{-5/2}$, where r_c is the core radius. Even for such a distribution, identical point masses will undergo core collapse within a time (see discussion in Binney & Tremaine 1987)

$$t_{\text{cc}} \approx 16 t_{\text{rlx,h}} \quad (3)$$

where $t_{\text{rlx,h}}$ is the relaxation time at the half-mass radius. Note that this is a factor ~ 20 times less than the time needed for the cluster to evaporate (Binney & Tremaine 1987). Core collapse of single objects will formally lead to infinite density at the center. In globular clusters and similar systems, this is avoided by the intervention of binaries: three-body and four-body scattering can transfer energy from binaries to the stellar velocity dispersion, heating the cluster and stabilising the density at the center (see Gao et al. 1991; Fregeau et al. 2003; Giersz & Spurzem 2000 for cluster simulations involving primordial binaries). As we now show, however, when the velocity dispersion is high enough (as it is in many observed galactic nuclei), binaries cannot prevent core collapse.

3 THE INSUFFICIENCY OF BINARIES

As shown first by Heggie (1975), binary-single interactions tend to harden hard binaries, and soften soft binaries. Only hardening will inject energy into the cluster and slow core collapse, hence we only need to consider hard binaries. For equal-mass objects the hard/soft boundary is approximately

where the orbital energy per object is equal to the kinetic energy of field stars (Quinlan 1996a). Suppose that the stellar velocity dispersion is v_{res} at the resolution radius r_{res} for a particular galactic nucleus. Then at the hard/soft boundary the semimajor axis a is given by

$$2Gm/a \approx v_{\text{res}}^2. \quad (4)$$

Any binary emits gravitational radiation as it orbits. If the time for the binary to merge by gravitational wave emission is less than the time for the binary to interact with field stars, then the binary does not heat the cluster. For a fixed semimajor axis, the merger time is maximised for a circular orbit, so we assume $e = 0$ to be conservative. For comparison, if $e \approx 0.7$ (the mean for a thermal distribution), the merger time is decreased by a factor ~ 10 for fixed a . The rate of change in the semimajor axis from gravitational radiation, and corresponding merger time for a circular orbit, is then (Peters 1964)

$$\begin{aligned} da/dt &= -\frac{64}{5} G^3 \mu m_{\text{bin}}^2 / (c^5 a^3) \\ \tau_{\text{merge}} = a/|da/dt| &= \frac{5}{128} c^5 a^4 / (G^3 m^3) \\ &= \frac{5}{8} (c/v_{\text{res}})^5 (Gm/v_{\text{res}}^3) \end{aligned} \quad (5)$$

where $m_{\text{bin}} = m_1 + m_2$ is the total mass of the binary, $\mu = m_1 m_2 / m_{\text{bin}}$ is the reduced mass, in the second line we assume $m_1 = m_2 = m$, and in the third line we substitute $a = 2Gm/v_{\text{res}}^2$.

The timescale for a three-body interaction is $\tau_{3\text{-bod}} = 1/(n\Sigma v)$, where n is the number density, $v \approx \sqrt{2}v_{\text{res}}$ is the relative speed, and $\Sigma = \pi r_p^2 [1 + 2G(m_{\text{bin}} + m)/(r_p v_{\text{res}}^2)]$ is the interaction cross section, where r_p is the distance of closest approach. For $r_p \approx a$ and three equal masses, a binary at the hard/soft boundary has $\Sigma \approx 4\pi a^2 \approx 16\pi G^2 m^2 / v_{\text{res}}^4$. Substituting $n = \rho/m$ we find

$$\tau_{3\text{-bod}} \approx v_{\text{res}}^3 / (16\sqrt{2}\pi\rho G^2 m). \quad (6)$$

The ratio between the merger and three-body timescales is then

$$\tau_{\text{merge}}/\tau_{3\text{-bod}} \approx 44(c/v_{\text{res}})^5 G^3 \rho m^2 / v_{\text{res}}^6. \quad (7)$$

This ratio needs to exceed unity for the typical binary to interact before it merges. Using the average density $\rho \approx \bar{\rho} = M/(4\pi R^3/3)$ and assuming a roughly constant velocity dispersion $v_{\text{res}}^2 = GM/R$, we find after some manipulation that $\tau_{\text{merge}}/\tau_{3\text{-bod}} > 1$ implies

$$\begin{aligned} m &\gtrsim \frac{1}{3} (v_{\text{res}}/c)^{5/2} M \\ &\approx 20 M_\odot v_{\text{res},3}^{5/2} M_8 \end{aligned} \quad (8)$$

where $v_{\text{res}} = 10^3 v_{\text{res},3} \text{ km s}^{-1}$. A cluster made of any point masses lighter than this cannot support itself by binary heating.

A loophole might appear to be that when there is bulk rotation (and hence a reduced velocity dispersion) or a density profile in which the relative speed at the center is much less than $(GM/R)^{1/2}$, binaries wide enough not to merge quickly could still heat the distribution. However, suppose that a binary has tightened by interactions to the point that $a = 2Gm/v_{\text{res}}^2$, as considered above. Its specific binding energy is then $G\mu/(2a) = Gm/(4a) = GM/(8R)$, because $v_{\text{res}}^2 = GM/R$ and thus $a = 2R(m/M)$. Even if the cluster is 100% binaries, the total binding energy liberated by hardening is therefore $GM^2/(8R)$. The minimum binding

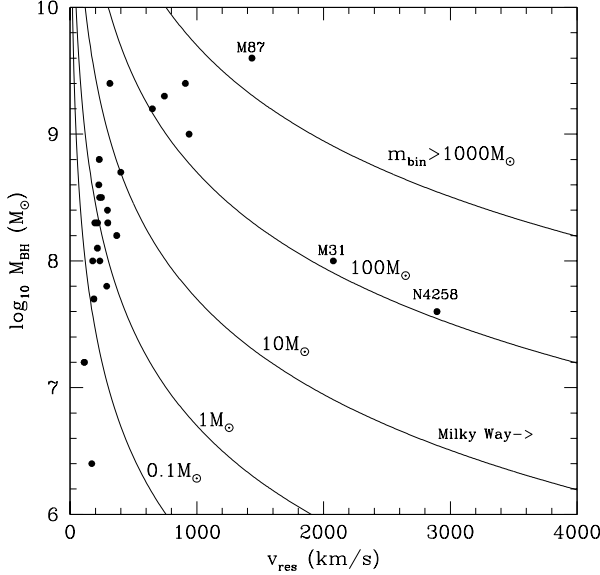


Figure 1. Inferred black hole masses and stellar speeds at resolution radius, derived from Table II of (Ferrarese & Ford 2005) with updates for M31 (Bender et al. 2005), and the Milky Way (Ghez et al. 2005), which is far to the right of the diagram at $v_{\text{res}} = 1.2 \times 10^4 \text{ km s}^{-1}$. The curved lines are labeled by the minimum mass of identical point masses such that if they make up the dark mass, binaries can in principle heat the system and delay core collapse. Several galaxies have $M_{\text{min}} > 100 M_{\odot}$ (the Galaxy has $M_{\text{min}} \approx 400 M_{\odot}$) and hence no reasonable stellar component could heat the system.

energy of a cluster with mass M and outer radius R is obtained when all the mass is in a thin spherical shell at radius R , in which case the binding energy is $GM^2/(4R)$. Even in this case, therefore, the maximum effect of binaries (prior to their reaching the previously considered semimajor axis $a = 2Gm/v_{\text{res}}^2$) is to increase the cluster binding energy, and hence the cluster radius, by 50%. A smaller binary fraction, a more concentrated cluster, or nonzero eccentricities for the binaries will all reduce this number. Therefore, if binaries that are hard relative to v_{res} merge quickly by gravitational radiation, no possible configuration of velocities or densities can allow binaries to stall collapse significantly.

Figure 1 plots the black hole mass versus the stellar velocity at the resolution radius, along with the minimum mass of point masses that would allow binary heating. Several galactic nuclei cannot be heated by masses lower than $100 M_{\odot}$, including the Galaxy, M87, M31, and NGC 4258. If such masses were assembled, the number of objects would therefore be small for a given dark mass, which would reduce the relaxation time dramatically (see equation 1) and would mean that the evaporation time $t_{\text{evap}} \approx 300 t_{\text{rlx}}$ would be much less than a Hubble time. Therefore, even with an implausible collection of $> 100 M_{\odot}$ objects in binaries, the cluster would still disintegrate rapidly.

4 CORE COLLAPSE

If core collapse happens, what is the result? Cohn (1980) found that the density profile approaches $n \propto r^{-2.23}$. For

ease of calculation, and to be conservative, we will assume a shallower profile $n \propto r^{-2}$, which is appropriate for a singular isothermal sphere. In such a profile, the total mass interior to radius r is proportional to r , and the velocity dispersion is constant with radius.

In the high density central regions, even point masses can merge because they emit gravitational radiation. Quinlan & Shapiro (1989) showed that for a relative speed v at infinity between two masses with reduced mass μ and total mass m_{tot} , there will be a mutual capture if the pericenter distance of approach r_p satisfies

$$r_p < r_{p,\text{max}} = \left(\frac{85\pi\sqrt{2}}{12c^5} \right)^{2/7} G\mu^{2/7} m_{\text{tot}}^{5/7} v^{-4/7}. \quad (9)$$

For equal masses and $v \approx \sqrt{2}v_{\text{res}}$, the cross section for merging in the gravitationally focused limit is

$$\Sigma_{\text{merge}} \approx 2\pi r_{p,\text{max}} (Gm_{\text{tot}}/v^2) \approx 19 \left(\frac{Gm}{c^2} \right)^2 \left(\frac{c}{v_{\text{res}}} \right)^{18/7}. \quad (10)$$

Over a time T , the probability of merger of an average point mass is then $P = Tn\Sigma_{\text{merge}}$. The average number density is $\bar{n} = (M/m)/(4\pi R^3/3)$. At this density, we find after some algebra that the probability is

$$\bar{P} \approx 4T \frac{m}{M} \left(\frac{v_{\text{res}}}{c} \right)^{10/7} \frac{v_{\text{res}}^3}{GM}. \quad (11)$$

With the rough approximation that $n \approx \bar{n}(r/R)^{-2}$ and $M(< r) \approx (r/R)M$, this implies that the enclosed mass M_{merge} inside of which the masses merge in time $T = 10^9 T_9$ Gyr is

$$\begin{aligned} M_{\text{merge}} &\approx \bar{P}^{1/2} M \approx 3(c^3/G)^{1/2} T^{1/2} m^{1/2} (v_{\text{res}}/c)^{31/14} \\ &\approx 5 \times 10^5 M_{\odot} T_9^{1/2} (m/1 M_{\odot})^{1/2} v_{\text{res},3}^{31/14}, \end{aligned} \quad (12)$$

or just M if $\bar{P} > 1$. The net result is that even for low-mass objects, core collapse will lead to the formation of a large single mass at the center of the distribution. However, as is clear from Equation (1), if the component masses are small enough then the relaxation time is so large that core collapse will not occur. We now address this situation.

5 DYNAMICAL FRICTION AND STELLAR-MASS BLACK HOLES

Suppose that the particles comprising the matter are very low-mass indeed, such as elementary particles. Suppose also that, like hypothesised dark matter, the particles interact neither with themselves nor with ordinary baryonic matter in any way but gravitationally. If in some improbable circumstance the particles have collected in a cluster of total mass M and radius R , what will affect them?

Because the particles have low mass, any more massive objects that enter their region will sink to the center via dynamical friction. The characteristic time for a mass m to sink is (see Binney & Tremaine 1987 for a discussion)

$$\tau_{DF} \approx v_M^3 / [4\pi\xi \ln \Lambda G^2 \rho m] \quad (13)$$

where v_M is the speed of the massive object, $\ln \Lambda$ is a Coulomb logarithm, and

$$\xi = \text{erf}(X) - \frac{2X}{\sqrt{\pi}} e^{-X^2} \quad (14)$$

with $X \equiv v_M/(\sqrt{2}\sigma)$. If $v_M \approx v_{\text{res}}$, then $\xi \approx 0.2$. Adopting as before $v_{\text{res}} \approx (GM/R)^{1/2}$ we find

$$\begin{aligned} \tau_{DF} &\sim 0.2(M/m)(GM/v_{\text{res}}^3) \\ &\approx 8 \times 10^9 \text{ yr } M_8^2 (1 M_\odot/m) v_{\text{res},3}^{-3}. \end{aligned} \quad (15)$$

This implies that for systems such as the central region of M31, where $v_{\text{res}} \approx 2000 \text{ km s}^{-1}$ and $M > 10^8 M_\odot$ (Bender et al. 2005), even ordinary stars will sink to the center of the mass distribution within a few Gyr, or much less if the dark matter is more concentrated. Therefore, all ordinary stellar processes that would proceed around a supermassive black hole will also proceed around a concentrated region of noninteracting particles, except that stars inside the region will sink to the center rapidly (see Quinlan 1996b). Thus if the dynamical friction time at the average density $\bar{\rho} = M/(4\pi R^3/3)$ is less than a few Gyr, stars and compact objects that enter the region will collide, merge, and have prime conditions for forming a large single mass.

The rate of interactions of stars with the central concentrated region is less for smaller regions. Suppose that the non-stellar matter is very concentrated, say with a radius just a few times the radius of a black hole with the same mass. Then, the arguments used to estimate rates of extreme mass ratio inspirals also apply here. These arguments suggest that stellar-mass black holes will spiral into supermassive black holes at a rate not less than $\sim 10^{-8} \text{ yr}^{-1}$ (Hils & Bender 1995; Sigurdsson & Rees 1997; Miralda-Escudé & Gould 2000; Freitag 2001, 2003; Ivanov 2002; Hopman & Alexander 2005). Therefore, regardless of how compactly the dark matter is distributed, if stellar-mass black holes exist they will enter the mass distribution in much less than a Hubble time.

The mass accreted by a black hole during inspiral is comparatively small. For example, consider a constant-density region $\rho = \bar{\rho} = M/(\frac{4}{3}\pi R^3)$ with nonrelativistic particles moving at an average speed $v_{\text{res}} = (GM/R)^{1/2}$ relative to the black hole. The cross section for absorption by a black hole of mass m is $\Sigma = (4Gm/c^2)(2Gm/v_{\text{res}}^2)$, so during a time τ_{DF} the black hole will accrete a mass

$$\begin{aligned} \Delta m &= \rho \Sigma v_{\text{res}} \tau_{DF} \\ &\approx 0.4 (v_{\text{res}}/c)^2 m. \end{aligned} \quad (16)$$

This is therefore only a small fraction of the original mass. Similarly, if after inspiral the black hole is fixed at the center of the mass distribution, it accretes little mass.

This conclusion changes if the black hole wanders freely around the dark matter distribution. This could happen if, for example, multiple massive objects enter the dark matter region and scatter each other frequently. In this case, for the same assumptions as before, the mass accretion rate $\dot{m} = \rho \Sigma v_{\text{res}}$ becomes

$$\dot{m} = 2 \frac{m^2}{M^2} \frac{\sigma^5}{G c^2}, \quad (17)$$

implying a growth time

$$\begin{aligned} T_{\text{growth}} &= \frac{1}{2} (M/m) (G M c^2 / \sigma^5) \\ &\approx 2 \times 10^{14} \text{ yr } M_8^2 (m/10 M_\odot)^{-1} v_{\text{res},3}^{-5}. \end{aligned} \quad (18)$$

This is not constraining on most supermassive black hole candidates, but for the Galaxy ($M \approx 4 \times 10^6 M_\odot$; Ghez et al. 2005) and $v_{\text{res}} \approx 1.2 \times 10^4 \text{ km s}^{-1}$ within 45 AU (Ghez et al. 2005), the growth time is $T_{\text{growth}} \sim 4 \times 10^5 \text{ yr}$.

Radio observations (Reid & Brunthaler 2004; Shen et al. 2005) suggest that at least $4 \times 10^5 M_\odot$ is contained within $\sim 0.5 \text{ AU}$ of the position of Sgr A*, which lowers the accretion time to at most $\sim 100 \text{ yr}$. Note that a doubling of mass decreases the time to the next doubling by a factor of two, so substantial growth results in a runaway. Note also that because by assumption the only matter entering the black hole does so with prompt infall and without release of radiation, the growth is not limited by the Eddington rate. If the particles are baryonic or otherwise have a reasonable strength of self-interaction, then the accretion rate is greatly enhanced, up to a possible Eddington-like maximum.

Finally, suppose that the non-luminous matter is in fact in the form of a star supported by pressure gradients rather than by simple motion as we have assumed up to this point. An example would be fermion balls (see, e.g., Tsiklauri & Viollier 1998), which are collections of massive neutrinos supported by degeneracy pressure. In that case, clearly a stellar-mass black hole captured by the star will consume matter at the star's center and remove pressure support, leading to rapid destruction of the star.

6 CONCLUSIONS

In the past decade, thanks to many observational developments, the case for supermassive black holes in the centers of many galaxies has gone from strong to essentially inescapable. We have shown that for many specific galactic nuclei, the observational constraints are strong enough to rule out binary heating, hence the relevant evolution time is the time to core collapse. This is a factor of ~ 20 less than the time to evaporation, which has previously been used as the conservative standard for stellar cluster persistence. For many individual galactic nuclei, therefore, the combination of time to core collapse and lack of binary heating rules out dense stellar clusters as an alternate explanation for the inferred dark mass. Specifically, the Galaxy, NGC 4208, and M31 have core collapse times $< 2 \text{ Gyr}$ for $0.5 M_\odot$ objects and cannot be stabilised by binaries less than $100 M_\odot$. M32 also has a core collapse time $< 2 \text{ Gyr}$, but could in principle be stabilised by stellar-mass binaries. All other sources currently have core collapse times $> 200 \text{ Gyr}$ for $0.5 M_\odot$ objects.

The only remaining possibilities are concentrated regions of noninteracting low-mass particles or self-supported exotic objects such as a fermion balls (Tsiklauri & Viollier 1998). Even in this case, we have shown that dynamical evolution of the stars and black holes near the centers of galaxies will cause multiple stellar-mass black holes to fall to the center of the potential, if black holes exist at all. Such black holes would consume any high-mass exotic pressure-supported objects, and would also accrete a noninteracting cluster of particles if allowed to move around freely. Therefore, the existence of stellar-mass black holes would lead to the production of supermassive black holes in many specific sources even if the supermassive holes did not form in other ways. When combined with the high redshifts inferred from Fe K α lines in some Seyfert galaxies (Reynolds & Nowak 2003), dramatic deviations from standard physics are required to explain observations in ways not involving black holes.

ACKNOWLEDGMENTS

We thank Laura Ferrarese, Marc Freitag, and Doug Richstone for valuable discussions. This work was supported in part by NASA ATP grant NAG 5-13229, and by a senior NRC fellowship during a sabbatical at Goddard Space Flight Center.

REFERENCES

- Bender R. et al., 2005, ApJ, 631, 280
Binney J., Tremaine S., 1987, Galactic Dynamics (Princeton University Press)
Cohn H., 1980, ApJ, 242, 765
Ferrarese L., Ford H., 2005, SSRv, 116, 523
Fregeau J.M., Gürkan M.A., Joshi K.J., Rasio F.A., 2003, ApJ 593, 772
Freitag M., 2001, Classical and Quantum Gravity, 18, 4033
Freitag M., 2003, ApJ Lett., 583, L21
Gao B., Goodman J., Cohn H., Murphy B., 1991, ApJ, 370, 567
Ghez A.M., et al. 2005, ApJ, 620, 744
Giersz M., Spurzem R., 2000, MNRAS, 317, 581
Heggie D.C., 1975, MNRAS, 173, 729
Hils D., Bender P.L., 1995, ApJ Lett., 445, L7
Hopman C., Alexander T., 2005, ApJ, 629, 362
Ivanov P.B., 2002, MNRAS, 336, 373
Kim E., Lee H.M., Spurzem R., 2004, MNRAS, 351, 220
Maoz E., 1998, ApJ, 494, L181
Miralda-Escudé J., Gould A., 2000, ApJ, 545, 847
Peters P.C., 1964, Phys. Rev. B, 136, 1224
Quinlan G.D., 1996a, New Astr., 1, 35
Quinlan G.D., 1996b, New Astr., 1, 255
Quinlan G.D., Shapiro S.L., 1989, ApJ, 343, 725
Reid M.J., Brunthaler A., 2004, ApJ, 616, 872
Reynolds C. S., Nowak M. A., 2003, Phys. Rep., 377, 389
Shen Z.-Q., Lo K.Y., Liang M.-C., Ho P.T.P., Zhao, J.-H., 2005, Nature, 438, 62
Sigurdsson S., Rees M.J., 1997, MNRAS, 284, 318
Spitzer L., 1987, Dynamical evolution of globular clusters (Princeton University Press)
Tsiklauri D., Viollier R.D., 1998, ApJ, 500, 591